



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 6, 2026 – 06:32 PM UTC

PDB ID : 5SMK / pdb_00005smk
Title : PanDDA analysis group deposition of ground-state model of SARS-CoV-2 NSP14
Authors : Imprachim, N.; Yosaatmadja, Y.; von-Delft, F.; Bountra, C.; Gileadi, O.; Newman, J.A.
Deposited on : 2022-03-08
Resolution : 1.65 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

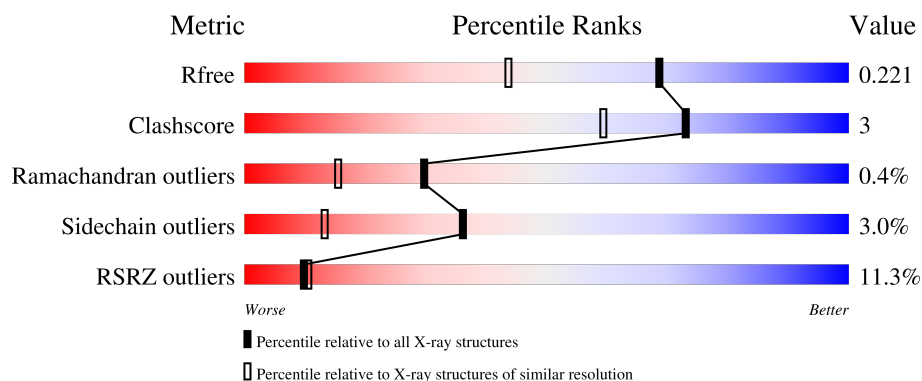
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.65 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	2563 (1.66-1.66)
Clashscore	190562	2662 (1.66-1.66)
Ramachandran outliers	187476	2621 (1.66-1.66)
Sidechain outliers	187428	2621 (1.66-1.66)
RSRZ outliers	180081	2564 (1.66-1.66)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	D	523	10% 79% 8% 12%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4116 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Proofreading exoribonuclease nsp14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	D	459	3672	2363	619	654	36	0	1	0

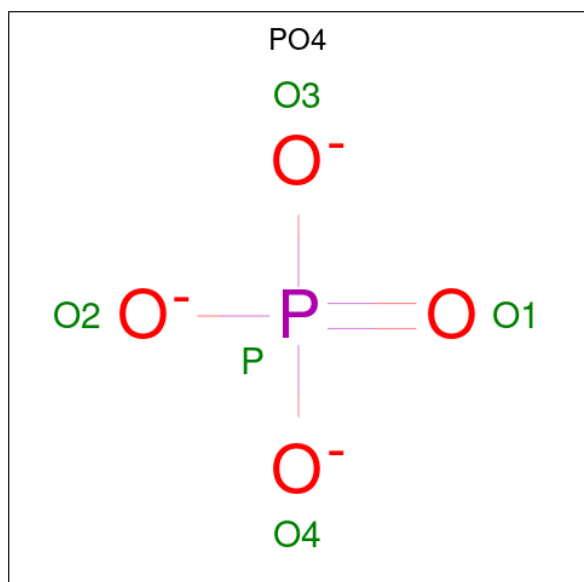
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	5	SER	-	expression tag	UNP P0DTD1
D	6	MET	-	expression tag	UNP P0DTD1

- Molecule 2 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	3	Total	Zn	0	0
			3	3		

- Molecule 3 is PHOSPHATE ION (CCD ID: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	D	1	Total	O	P	0	0
			5	4	1		
3	D	1	Total	O	P	0	0
			5	4	1		

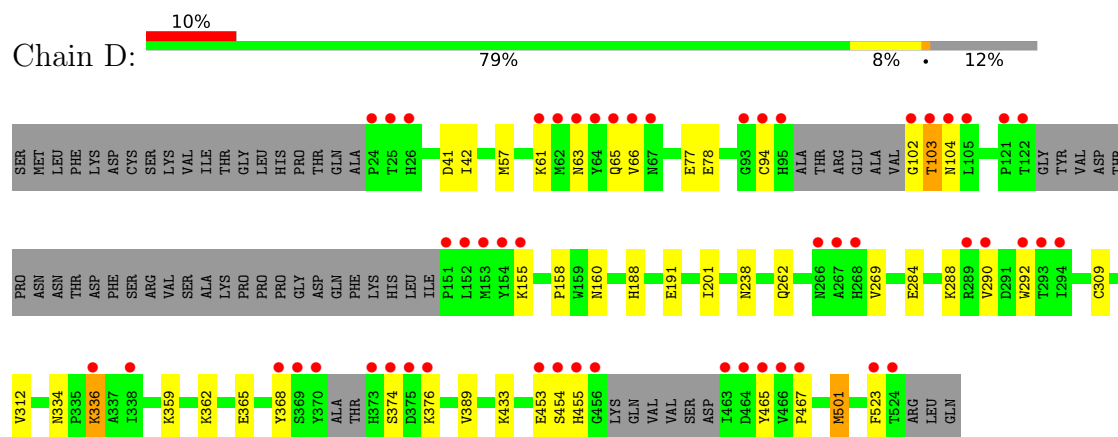
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	431	Total	O	0	0
			431	431		

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Proofreading exoribonuclease nsp14



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	67.54Å 68.09Å 138.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	69.31 – 1.65 69.21 – 1.66	Depositor EDS
% Data completeness (in resolution range)	50.9 (69.31-1.65) 50.9 (69.21-1.66)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.32 (at 1.65Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.191 , 0.217 0.197 , 0.221	Depositor DCC
R_{free} test set	3774 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å ²)	20.0	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 31.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.009 for k,h,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	4116	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.86% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, PO4

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	D	1.03	0/3775	1.24	3/5125 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	D	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	D	155	LYS	CA-C-N	5.83	126.25	120.31
1	D	155	LYS	C-N-CA	5.83	126.25	120.31
1	D	66	VAL	CA-C-O	-5.39	117.94	122.63

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	368	TYR	Peptide

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within

the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	3672	0	3533	23	0
2	D	3	0	0	0	0
3	D	10	0	0	0	0
4	D	431	0	0	3	1
All	All	4116	0	3533	23	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

All (23) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:42:ILE:HD13	1:D:57:MET:CE	2.26	0.66
1:D:42:ILE:HD13	1:D:57:MET:HE1	1.78	0.65
1:D:78:GLU:OE1	4:D:701:HOH:O	2.15	0.64
1:D:309:CYS:HA	1:D:501:MET:HE3	1.83	0.60
1:D:309:CYS:HA	1:D:501:MET:CE	2.33	0.59
1:D:41:ASP:OD2	1:D:188:HIS:HD2	1.86	0.57
1:D:290:VAL:HG11	1:D:292:TRP:CZ2	2.43	0.53
1:D:465:TYR:CZ	1:D:467:PRO:HA	2.49	0.48
1:D:290:VAL:HG11	1:D:292:TRP:CE2	2.49	0.48
1:D:158:PRO:HB2	1:D:160:ASN:OD1	2.13	0.47
1:D:374:SER:C	1:D:376:LYS:H	2.23	0.46
1:D:433:LYS:NZ	4:D:713:HOH:O	2.40	0.46
1:D:102:GLY:C	1:D:104:ASN:H	2.24	0.45
1:D:77:GLU:H	1:D:77:GLU:CD	2.26	0.44
1:D:362:LYS:HB2	1:D:365:GLU:HG2	2.00	0.44
1:D:362:LYS:O	1:D:365:GLU:HG2	2.19	0.42
1:D:61:LYS:HG2	1:D:63:ASN:ND2	2.34	0.42
1:D:284:GLU:HA	1:D:288:LYS:HG2	2.02	0.41
1:D:201:ILE:HA	1:D:238:ASN:OD1	2.20	0.41
1:D:334:ASN:OD1	1:D:336:LYS:HB2	2.20	0.41
1:D:262:GLN:NE2	4:D:735:HOH:O	2.53	0.41
1:D:312:VAL:HB	1:D:501:MET:HE1	2.03	0.41
1:D:374:SER:C	1:D:376:LYS:N	2.80	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:912:HOH:O	4:D:1007:HOH:O[4_556]	1.90	0.30

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	D	450/523 (86%)	434 (96%)	14 (3%)	2 (0%)	30 15

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	455	HIS
1	D	103	THR

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	D	399/460 (87%)	387 (97%)	12 (3%)	36 13

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	65	GLN
1	D	94	CYS
1	D	103	THR
1	D	191	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	269	VAL
1	D	336	LYS
1	D	359	LYS
1	D	389	VAL
1	D	453	GLU
1	D	454	SER
1	D	501	MET
1	D	523	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	63	ASN
1	D	67	ASN
1	D	188	HIS
1	D	245	GLN
1	D	268	HIS
1	D	408	ASN
1	D	518	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 3 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The

Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
3	PO4	D	605	-	4,4,4	1.08	1 (25%)	6,6,6	0.52	0
3	PO4	D	604	-	4,4,4	1.11	1 (25%)	6,6,6	0.80	0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	605	PO4	P-O1	2.13	1.55	1.50
3	D	604	PO4	P-O1	2.09	1.55	1.50

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å ²)	Q<0.9
1	D	459/523 (87%)	0.50	52 (11%)	10 10	8, 20, 71, 120	1 (0%)

All (52) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	151	PRO	9.2
1	D	373	HIS	8.9
1	D	95	HIS	8.4
1	D	374	SER	8.1
1	D	64	TYR	7.9
1	D	102	GLY	7.7
1	D	463	ILE	7.7
1	D	24	PRO	7.6
1	D	370	TYR	7.6
1	D	152	LEU	6.9
1	D	66	VAL	6.5
1	D	465	TYR	6.3
1	D	466	VAL	6.0
1	D	524	THR	5.8
1	D	105	LEU	5.7
1	D	368	TYR	5.2
1	D	456	GLY	5.2
1	D	122	THR	5.2
1	D	25	THR	4.9
1	D	103	THR	4.7
1	D	455	HIS	4.3
1	D	153	MET	4.2
1	D	154	TYR	4.0
1	D	369	SER	3.9
1	D	375	ASP	3.9
1	D	267	ALA	3.8
1	D	65	GLN	3.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	104	ASN	3.7
1	D	94	CYS	3.6
1	D	454	SER	3.6
1	D	268	HIS	3.6
1	D	464	ASP	3.6
1	D	292	TRP	3.5
1	D	121	PRO	3.3
1	D	290	VAL	3.3
1	D	376	LYS	3.3
1	D	523	PHE	3.2
1	D	26	HIS	3.1
1	D	289	ARG	3.1
1	D	62	MET	3.0
1	D	61	LYS	2.9
1	D	93	GLY	2.9
1	D	155	LYS	2.9
1	D	338	ILE	2.8
1	D	453	GLU	2.7
1	D	63	ASN	2.7
1	D	293	THR	2.5
1	D	67	ASN	2.4
1	D	336	LYS	2.4
1	D	467	PRO	2.2
1	D	266	ASN	2.1
1	D	294	ILE	2.1

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
3	PO4	D	604	5/5	0.81	0.18	37,51,61,62	0
3	PO4	D	605	5/5	0.82	0.15	41,42,56,58	0
2	ZN	D	602	1/1	0.99	0.04	17,17,17,17	0
2	ZN	D	603	1/1	0.99	0.02	26,26,26,26	0
2	ZN	D	601	1/1	1.00	0.02	15,15,15,15	0

6.5 Other polymers [i](#)

There are no such residues in this entry.